Capability of electron identification for the CALET measurement.

Lorenzo Pacini
Institute of Applied Physics (IFAC-CNR), Via Madonna del Piano, 10, 50019, Italy
University of Florence, Via Sansone, 1 - 50019 Sesto, Fiorentino, Italy
National Institute for Nuclear Physics (INFN), Piazza dei Caprettari, 70 - 00186 Rome, Italy
E-mail: l.pacini@ifac.cnr.it

Yosui Akaike[*]
University of Maryland, Baltimore County and NASA Goddard Space Flight Center, 8800
Greenbelt Rd. Greenbelt, MD 20771, USA E-mail: yosui.akaike@nasa.gov

for the CALET collaboration

The CALorimetric Electron Telescope, CALET, was installed on the International Space Station in August 2015 and it has been collecting data of high energy cosmic rays since October 2015.

The primary purpose of CALET mission is obtaining high-precision direct measurements of the electron+positron spectrum up to the multi-TeV energy region, which can provide unique information on the presence of nearby astrophysical sources and possible signals from dark matter. Other important objectives are the observation of nuclei spectra from proton to Fe up to several hundred TeV energies and the detection of gamma-ray emissions up to ~ 10 TeV with high-accuracy energy measurement.

The CALET instrument consists of a charge detection device composed of two layers of plastic scintillators, a finely-segmented sampling calorimeter (IMC) with imaging capabilities made of scintillating fibers (with total depth of 3 radiation lengths) and a homogeneous calorimeter (TASC) with PWO scintillating logs (27 radiation lengths).

Since protons are the largest background source for the electron measurements and the ratio of protons to electrons increases at higher energies, a proton rejection power of 10^5 is required in the TeV region. Owing to the thick and well-segmented calorimeter, CALET has capabilities to absorb most of shower energy for electrons in TeV region and to identify electrons from protons by the difference of shapes between electromagnetic showers and hadronic showers. The present work contains a comparative study, with different Monte Carlo simulation codes, of the proton rejection power obtained with optimized selection cuts. In order to maximize the rejection power we have also confirmed the consistency between flight data and Monte Carlo simulations for the variables obtained by the reconstruction of shower images and we have applied multivariate analysis (MVA) with boosted decision tree algorithms. The resulting capability of electron identification with CALET will be presented.

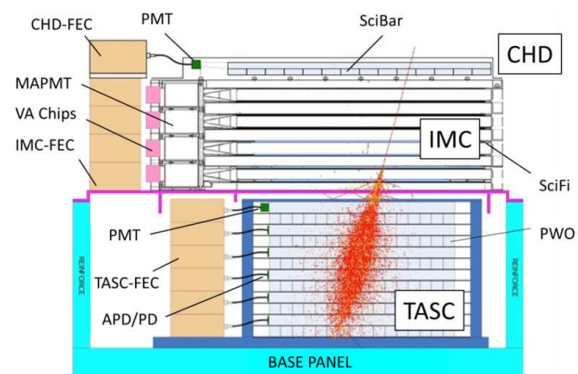[*]Speaker.

1. Introduction.

The CALorimetric Electron Telescope (CALET) is a calorimetric experiment for the direct measurement of high energy cosmic ray spectra. The apparatus was launched to the International Space Station (ISS) and installed on the Exposed Facility of the Japanese Experiment Module (JEM-EF) in August 2015. The cosmic ray observation started in October 2015 for an initial period of 2 years, with a maximum target period of 5 years. The CALET mission is designed for the measurement of the cosmic ray electron+positron (hereafter "electron") spectrum up to few TeV, the proton and nuclei (from Helium to Iron) spectra up to hundred Tev/nucleon and the precise measurement of the high energy gamma-rays. The direct measurements of cosmic ray spectra up to the "knee" region are crucial to investigate the acceleration and propagation of particles in the Galaxy. Electrons, during their diffusion in the Galaxy, lose their energy by synchrotron radiation and inverse Compton scattering so that TeV electrons observed at Earth position are likely originated in sources younger than 10^5 years and within 1 kpc distance from the Solar System; therefore the electron energy spectrum above 1 TeV can provide unique information about nearby cosmic ray sources, e.g. Supernova remnants. The measurement of high energy nucleus spectra provides important information about the propagation of particles in the Galaxy:for example the Carbon/Boron flux ratio is strongly dependent on the average thickness of interstellar matter which the cosmic rays traverse.

CALET also features a detector for the identification of Gamma Ray Burst signals, the Calet Gamma-ray Burst Monitor (CGBM), designed for the detection of photons from 7 keV to 20 MeV and consisting of two different kind of scintillators (LaBr$_3$ (Ce) and BGO).

2. The CALET calorimeter.

The CALET calorimeter is composed of three detectors (fig. 1): a *"Total Ab-Sorption Calorimeter"* (TASC), an *"IMaging Calorimeter"* (IMC) and a *"CHarge Detector"* (CHD).

The TASC is an homogeneous calorimeter with 12 layers oriented in X and Y directions: each layer is composed of 16 PWO scintillator crystal logs ($20 \times 19 \times 320$ mm^3). The total TASC thickness for vertical particle is 27 radiation lengths (X$_0$). The scintillation light produced in the top layer is read-out by Photo Multiplier Tubes (PMTs) and by an APD/PD pair for the crystals of the other layers, therefore each signal is amplified by two front-end circuits with different gains. The PMTs ensure a fast response in

Figure 1: *Representation of the CALET main detectors. An electromagnetic shower simulated with EPICS is also shown. The Z axis is the vertical direction.*

order to use these signals in the CALET trigger logic while the APD/PD and the double gain allow

a high dynamic range. The TASC design achieves a very good energy resolution for electrons, better than 2% at 1 TeV, and a good proton rejection power as discussed in the following sections. The IMC is a sampling calorimeter used for the detection of the early stage of the shower and the identification of the incoming particle direction. It consists of 16 layers of 448 scintillating fibers (SciFi) alternatively oriented in X and Y directions, interleaved with thin tungsten sheets for a total thickness of 3 X_0. The SciFis are coupled with 64-channel Multi-Anode PMTs. By using a Kalman filter algorithm[1] the expected angular resolution for high energy electrons is better than 0.5°. The CHD is composed of 2 layers, each with 14 plastic scintillator paddles oriented along X and Y directions respectively. Each paddle is read-out with a single PMT. The CHD can resolve individual elements from Z=1 to Z=40 with an excellent charge resolution.

3. Monte Carlo simulations.
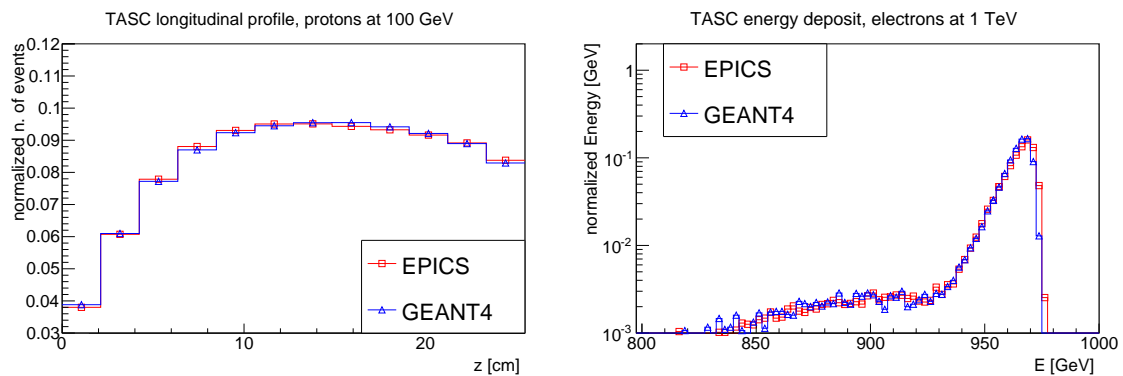
Figure 2: *Comparison between EPICS (red squares) and GEANT4 (blue triangles). Left: longitudinal profile of the proton shower at* 100 GeV *in the TASC. Right: the total energy deposited in TASC by electrons at* 1 TeV.

In order to study the CALET calorimeter performance, several Monte Carlo simulations of the detector were developed. The present work features the expected electron/proton (e/p) separation with simulations based on the EPICS[2] and GEANT4[3] packages. In order to validate the complex geometry implemented in the simulations, common benchmarks between EPICS and GEANT4 has been carried out. Protons and electrons with vertical incident trajectory are simulated at different energies (10 GeV, 100 GeV, 1 TeV and 10 TeV) by employing a plane generation surface at the top of the CHD. The comparison involved several variables like the total energy deposit in the CHD, IMC, and TASC and the longitudinal and lateral profile of the shower. A good agreement between the simulations was found, especially for the variables related to the TASC. Two examples are shown in fig. 2. At the same time, some discrepancies were found in the CHD and IMC total energy deposit for protons and electrons above 100 GeV, apparently due to different treatment of the backscatter particles; these are currently being investigated in detail to improve much more the capability of e/p separation at higher energies.

For the study of the e/p discrimination power, simulated particles are isotropically generated in order to cover a geometric region greater than the true CALET acceptance and take into account possible contamination from particles incident outside the acceptance and interacting with materials surrounding the detectors.

EPICS simulation   Electrons and protons are simulated with EPICS version 9.20 hadronic interaction model DPMJET-III, in the energy range from 2 GeV to 20 TeV. The incoming flux is generated as a single power law spectrum with spectral index equal to -1. In order to extend the simulated proton spectrum and avoid very long computation time, protons between 20 TeV and 1 PeV are simulated with spectral index of -2.5.

GEANT4 simulation   Protons and electrons are simulated with GEANT4, version 10.01.p02 and FTFP-BERT physics list, using a single power law spectrum with spectral index equal to -1. Electrons are simulated from 2 GeV to 20 TeV while protons are generated from 10 GeV to 100 TeV.

In order to calculate the expected proton contamination in the electron flux measurement, an event by event re-weighting procedure of the simulated spectra is applied. The electron simulations are weighted according to a single power law with spectral index -3.15 and a normalization factor of ~ 300: these parameters are obtained by fitting the electron flux measured by the AMS experiment[4] in the 30 GeV $-$ 1 TeV energy range. Proton simulations are weighted according to the AMS spectrum by using the function described in [5] and taking into account the conversion between the proton rigidity and the kinetic energy.

4. Event selection before the e/p discrimination cut.

The present analysis takes into account events in the so called CALET acceptance A (acc. A). A particle is in acc. A if its initial direction crosses the top surface of the CHD, the top and the bottom of the TASC but with a fiducial cut excluding a lateral strip of width equal to 1.9 cm. The geometrical factor of the acc. A is ≈ 416 cm^2 sr: this is smaller than the total CALET acceptance but has the advantage that for all the events the electromagnetic showers are almost fully contained in the calorimeter, hence only a small correction of the energy deposited in the calorimeter is needed in order to retrieve the true energy of the electron (a correction $< 5\%$ is needed for electrons of 1 TeV). In the present work all the energy bins are defined using the energy deposited in the calorimeter and applying this small correction factor.

Before the cut used for the e/p identification, some selections are needed in order to take into account the detector trigger and avoid contamination from events outside the acceptance and heavier nuclei.

The main CALET trigger, named the *"High Energy Trigger"* (HET), is optimized for the observation of high energy electrons, requiring a big energy deposit in the last 2 layers of the IMC (~ 25 MIP) and in the first layer of the TASC (~ 90 MIP). The first selection of the analysis is a software trigger (SW-HET): the channels involved in the SW-HET are the same in the hardware HET but the thresholds are slightly higher (100 MIP for the TASC and 40 MIP for the IMC). This cut is effective in order to avoid the need for an accurate simulation of the discriminator thresholds implemented in the electronics. The efficiency of the SW-HET cut is very high for electrons,

$> 95\%$ at 30 GeV and almost 100% above 100 GeV, while it is small for protons ($< 50\%$) below 1 TeV.

In the CALET analysis framework, the trajectory of the particle is reconstructed using several different algorithms mainly based on the IMC fiber signals: in the present work only the track reconstructed with the Kalman filter technique is taken into account. The algorithm efficiency is very high for both electrons and protons. All particles in acc. A are correctly reconstructed in such acceptance, with efficiency larger than 98%.

On the other hand, some events generated outside of acc A. can be reconstructed in acc. A. In order to avoid a considerable acceptance contamination, a selection named *"IMC Shower Concentration"* (IMC-SC) is added: exploiting the fact that the electromagnetic shower starts in the IMC tungsten layers, for an electron a large energy deposit near the true particle direction is present in the last layer of the IMC. The IMC-SC cut rejects events with the energy deposit in the last IMC X (Y) layer within 1 Moliere radius of the tungsten (equal to 9 SciFi) from the reconstructed particle direction below 40% of the energy deposit in the last IMC X (Y) layer. After this cut the expected contamination in the electron flux due to events initially outside acc. A is $< 1.5\%$.

The last applied selection is a charge cut in order to reject He and heavier nuclei: the event is rejected if the quadratic mean of the energy deposit in the paddles crossed by the particle reconstructed trajectory is > 3.5 MIP. Only a few He events survive this selection, while the efficiency for electrons is high ($\sim 97\%$ up to 1 TeV).

The total efficiency of these selection cuts is almost constant with the energy for electrons, with a value of $(97.5 \pm 0.2)\%$ at 100 GeV and $(97.6 \pm 0.2)\%$ at 1 TeV, while for protons is smaller and weakly depends on the energy ($(22.8 \pm 0.2)\%$ at 100 GeV and $(26.3 \pm 0.3)\%$ at 1TeV).

5. e/p separation with a single cut.
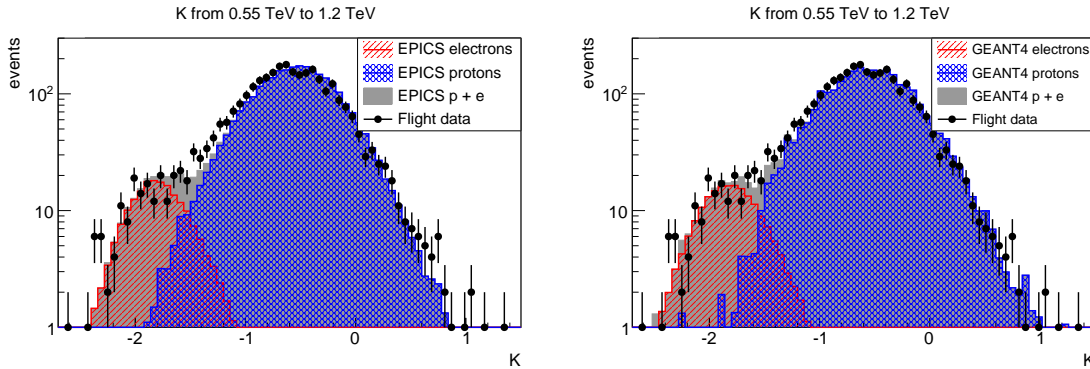
Figure 3: *Comparison between flight data and Monte Carlo simulations regarding the K variables in the energy range from 550 GeV and 1.21 TeV. Left: K distribution with the EPICS electron simulation (red histogram), proton simulation (blue histograms) and the flight data (black histogram). The filled gray histogram is the distribution obtained as the sum of the proton and electron simulations. Right: same comparison with GEANT4 simulations.*

Since the proton abundance in cosmic rays is 100-1000 times larger than the electron one, a very good proton rejection power is required. The event selection described in sec. 4 has a very high
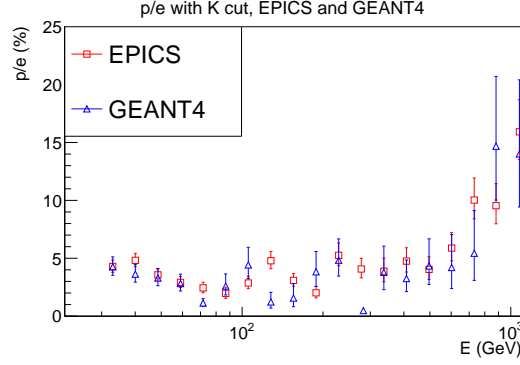
Figure 4: *Comparison of the expected residual proton contamination using the K variable e/p cut with EPICS (red squares) and GEANT4 (blue triangles) simulations in the energy bins from* 30 GeV *to* 1 TeV.

efficiency for electrons while rejecting $\sim 75\%$ of protons, but the proton contamination still remains considerable thus requiring additional strong rejection. The e/p discrimination cut discussed here involves the development of the shower in the TASC. Two variables are used: the first (F_E) is the energy deposited in the last (bottom) TASC layer divided by the total energy deposited in the TASC; this is related to the longitudinal shower development and in hadronic showers this value is bigger than in electromagnetic ones. The second variable (R_E) is the energy weighted spread in the first (topmost) TASC layer: it takes into account the lateral profile of the early part of the shower:

$$R_E = \sqrt{\frac{\sum_j (\Delta E_j \cdot (x_j - x_c)^2)}{\sum_j \Delta E_j}},$$

where x_c is the coordinate of the reconstructed particle track in the first TASC layer, x_j and ΔE_j are the coordinate of the center and the energy deposit in the jth PWO log respectively. These variables are chosen not only to achieve a high proton rejection power but also in order to get a good agreement between the Monte Carlo simulation and the flight data. In order to improve the cut efficiency for electrons, while keeping a good discrimination power, the cut is applied on a single variable (K) obtained by combination of the two variables as follow:

$$K = log_{10}(F_E) + 0.65 \cdot R_E$$

where the 0.65 numeric parameter is selected in order to maximize the rejection power at 1 TeV. A comparison of the distribution of the K variables calculated using the Monte Carlo simulations and the flight data above 550 GeV is shown in fig. 3. A reasonable agreement is achieved with both EPICS and GEANT4 in the energy range from 30 GeV and 1 TeV.

By employing the Monte Carlo simulation, the thresholds for this cut are chosen in order to achieve a constant efficiency of 70% for electrons in each energy bin. Fig. 4 shows the expected residual proton contamination as computed with the EPICS and GEANT4 simulations. The proton contamination is smaller than 5% up to 700 GeV and about of order of 15% at 1 TeV. This result shows also an overall agreement between the two different Monte Carlo simulations. Some small discrepancies are present, which will translate in differences in the electron flux measurement of at

most few percent, due to the proton contamination subtraction.

6. e/p separation with Multi Variate Analysis.
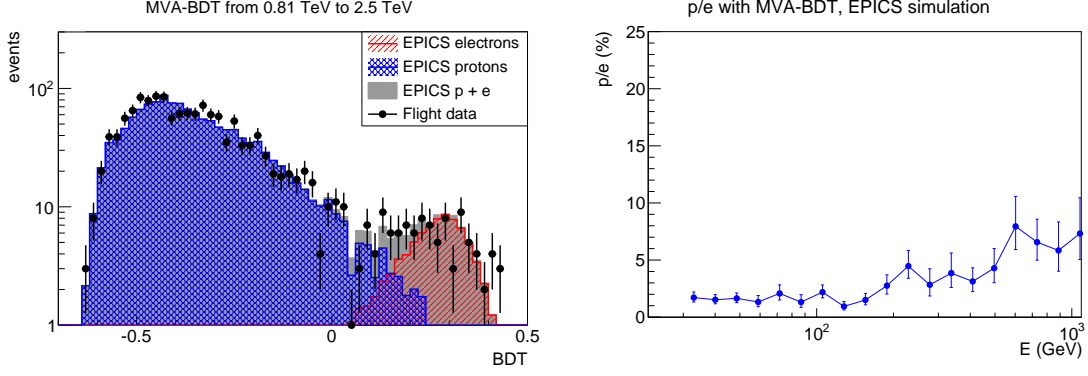
Figure 5: *Left: comparison between the flight data and the Monte Carlo simulations based on EPICS regarding the BDT variable in the energy range from* 810 GeV *and* 2.5 TeV. *Right: expected proton contamination in the electron energy bins from* 30 GeV *to* 1 TeV *computed with the EPICS simulations*

To improve the e/p discrimination up to the multi-TeV region, while keeping a high efficiency for electrons, a Multi Variate Analysis (MVA) approach has been developed. Preliminary comparison[6] between different MVA algorithms showed that the Boosted Decision Trees (BDT) is the most promising one. The MVA analysis has been developed and optimized with the EPICS simulation, while GEANT4 has not been used so far because the statistics of the high energy protons generated with the GEANT4 simulation is not enough for the training of the BDT algorithm. In addition to the F_E and R_E other discriminant variables are included.

The first one is the IMC-SC already described in sec. 4 where it was used as a weak cut in order to avoid a big acceptance contamination. This selection has a proton rejection power similar to the rejection power of the R_E variable if used as a strong cut (i.e. with lower efficiency for electrons).

Other variables represent the TASC longitudinal shower profile parameters: by using the energy deposit in each TASC layer, the longitudinal shower profile is fitted with a Gamma distribution:

$$\frac{dE}{dt} = E_0 \frac{t^{\alpha-1} e^{\frac{t}{\theta}}}{\Gamma(\alpha)\theta^\alpha}$$

where t is the depth in unit of X_0, θ is a scale parameter (in unit of X_0), α is a shape parameter and E_0 is the total energy of the shower. Three variables related to this fit are included in the BDT algorithm: point of the maximum of the function $t_{max} = (\alpha - 1)\theta$, the θ parameter and the χ^2 of the fit.

Similarly, the longitudinal signal distribution in the IMC can be described with $dE/dt = p_0 + p_1 \cdot t^2$ and both the parameter p_0, p_1 and the χ^2 of the event fit are used in the BDT analysis.

The generated proton and electron events are separated in two samples, the first one used for the BDT training, and the other one for the calculation of the efficiency and the proton contamination:

6

for each energy bin an independent BDT training is carried out in order to optimize the rejection power.

In each energy bin only the variables that show a good agreement between the EPICS Monte Carlo simulations and the flight data are included in the BDT algorithm: above 500 GeV all the variables but the IMC-SC is used.

In the left plot in fig. 5 the distribution of the BDT variable for simulated electrons and protons (test samples) are compared with the flight data in the region above 810 GeV. A good separation between the proton and electron peaks is visible and the Monte Carlo simulations show a reasonable agreement with the flight data. The expected proton contamination corresponding to a constant 70% efficiency of the rejection cut, shown on the right side of fig. 5, is smaller than the one computed with the K cut analysis: it remains below 10% up to 1 TeV.

7. Conclusions.

The analysis discussed here shows a reasonable agreement between the expected proton contamination in the measured electron spectrum computed with the EPICS and GEANT4 simulations with the K cut analysis. This simple approach achieves a small residual proton contamination in the low energy bins while this becomes higher at 1 TeV. In order to improve the proton rejection power a MVA approach has been also developed: with the same efficiency for the electron selection the expected proton contamination is smaller than the one expected with the K cut and remains below 10% up to 1 TeV.

These simulation results are validated with precise comparison with the flight data: a very similar analysis procedure is discussed in [7] and applied for the preliminary measurement of the CALET electron flux up to 1 TeV.

References

[1] P. Brogi et al. (CALET Collaboration), *CALET measurements with cosmic nuclei: expected performances of tracking and charge identification* PoS(ICRC2015)595

[2] K. Kasahara, Proc. of 24th International Cosmic Ray Conference (Rome, Italy), Vol. 1 (1995) 399; "EPICS Home Page" http://cosmos.n.kanagawa-u.ac.jp/EPICSHome/.

[3] S. Agostinelli et al. *Geant4 a simulation toolkit* Nuclear Instruments and Methods in Physics Research Section A, Volume 506, Issue 3, 1 July 2003, Pages 250-303

[4] M. Aguilar et al. (AMS Collaboration), *Electron and Positron Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station* Phys. Rev. Lett. 113, 121102, (2014)

[5] M. Aguilar et al. (AMS Collaboration), *Precision Measurement of the Proton Flux in Primary Cosmic Rays from Rigidity 1 GV to 1.8 TV with the Alpha Magnetic Spectrometer on the International Space Station.* Phys. Rev. Lett. 114, 171103, (2015)

[6] F. Palma et al.(CALET Collaboration), *Simulation studies of the expected proton rejection capabilities of CALET* PoS(ICRC2015)1196.

[7] Yoichi Asaoka et al.(CALET Collaboration), *Analysis and Preliminary Results for the Cosmic Ray Electron Spectrum from CALET* PoS(ICRC2017) 205